Journal of Aeronautics, Astronautics and Aviation, 2014, vol. 46, iss. 2, pp114-123
DOI:10.6125/14-0324-789

The Application of Aeronautical Decision-making Support Systems for Improving Pilots' Performance in Flight Operations

Wen-Chin Li[1], Lun-Wen Li[2], Don Harris[3], Yueh-Ling Hsu[4]

[1]Safety and Accident Investigation Center, Air Transport Department, Cranfield, Bedford, MK430AB, Cranfield University, United Kingdom

[2]Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan

No. 101, Section 2, Kuang-Fu Road, Hsinchu, Taiwan 30013, R.O.C.

[3]Engineering and Computing Department, Coventry University, Priory Street, Coventry, CV1 5FB, United Kingdom

[4]Department of Air Transportation, Kainan University, Taiwan

No.1 Kainan Road, Luzhu Shiang, Taoyuan 33857, Taiwan

Abstract

Operating a high-technology commercial airliner is not only an issue in psychomotor skill performance but also of a real-time decision-making involving situation awareness and risk management within a limited-time condition. The number of aircraft accidents attributable solely to mechanical failures has decreased markedly in recent years, but the contribution of human error has declined at a much slower rate. Previous research demonstrated a belief rule-based decision support system has provided more reliable and informative performance after training. The purpose of this research was to identify the best mnemonic-based method of decision support systems for improving commercial pilot's performance in the advanced cockpit. A total of 157 airline pilots, all qualified on the Boeing 747-400 evaluated the suitability of four different ADM methods: SHOR (Stimuli, Hypotheses, Options, Response); PASS (Problem identification, Acquire information, Survey strategy, Select strategy); FORDEC (Facts, Options, Risks & Benefits, Decision, Execution, Check); and DESIDE (Detect, Estimate, Set safety objectives, Identify, Do, Evaluate). Each was evaluated for six different types of decisions: go/no go; recognition-primed; response selection; resource management; non-diagnostic procedural; and creative problem-solving. Pilots regarded the FORDEC methodology as being the best in all decision-making scenarios, irrespective of the time available to make the decision. It was also rated as the best ADM method for promoting crew coordination. However, it was advised that practicing the FORDEC mnemonic in flight simulator was important before attempting to apply it in a real life situation.

***Keywords*:** Aeronautical Decision-making; Human Factors in Aviation; Mnemonic-based Training; Situational Awareness

Ⅰ. Introduction

Operating a high-technology commercial airliner is not only an issue in technical skill performance but also of a real-time decision-making involving situation awareness and risk management within a limited-time condition. The number of aircraft accidents attributable solely to mechanical failures has decreased markedly in recent years, but the contribution of human error has declined at a much slower rate (Shappell & Wiegmann, 2004). Figures vary but Jensen and Benel (1977) observed that decision errors contributed to 52% of all fatal general aviation accidents in the United States between 1970 and 1974. Diehl (1991) proposed that decision errors contributed to 56% of airline accidents and 53% of military accidents. More recently an analysis of military aviation accidents using the Human Factors Analysis and Classification System – HFACS (Wiegm ann & Shappell, 2003) found that 'decision errors' had the second highest rate of occurrence for being implicated as a causal factors (42.6%). In this study the decision errors included selecting inappropriate strategies to perform a mission; improper in-flight planning; making an inappropriate decision to abort a take-off or landing; or using improper remedial actions in an emergency. There is a raising need for developing aeronautical decision-making (ADM) training syllabus to improve aviation safety.

Shappell and Wiegmann (2004) using the same analytical framework found similar figures. They observed that decision errors contributed to 45% of accidents in the USAF and 55% in the US Navy. A similar study of airline accidents commercial in Taiwan using the HFACS taxonomy found that inadequate decision making was implicated in 70% of cases (Li, Harris, & Yu, 2008). In short, poor decision making is a serious threat to flight safety. In O'Hare's (2003) review of aeronautical decision-making he came to the conclusion that 'it is difficult to think of any single topic that is more central to the question of effective human performance in aviation than that of decision-making'.

Decision support systems include knowledge-based systems which integrates software and hardware to help decision maker compile useful information from a combination of raw data, standard operation procedurals and knowledge to identify and solve ill-defined problems (Power, 2002). There were lots of applications for decision support systems in different domains, such as aviation, medical healthcare, military command and control, and nuclear power station. Dreiseitl and Binder (2005) proposed that it is paramount importance for providing the quality assurance and validation of decision support systems in medical environment. Moreover, Kong, Xu, Body, Yang, Mackway-Jones and Carley (2012) suggested that a belief rule-based decision support system has provided more reliable and informative diagnosis recommendations than manual diagnosis using traditional rules during clinical uncertainties, and the diagnostic performance of the system can be significantly improved after training. Liao (2000) demonstrated there were great advantages for incorporate case-based reasoning and decision support system for training military officers the standard operation procedure (SOPs) in the decision-making process of command and control in the battlefield.

Ⅱ. The Definition of Aeronautical Decision-making

Decision making performance in the aviation domain is a joint function of the features of the tasks and the pilots' knowledge and experience relevant to those tasks. Aeronautical decision-making (ADM) is defined by the FAA (1991) as 'a

systematic approach to the mental process used by aircraft pilots to consistently determine the best course of action in response to a given set of circumstances'. In non-normal, abnormal and emergency situations pilots are often required to solve unexpected and ill-defined problems with only partial information available and while under time pressure. In flight decision making involves situation assessment, choice among alternatives and risk management (Orasanu & Connolly, 1993). Decisions are often made with ambiguous information available and under time pressure. McKinney (1993) suggested that stress and task saturation may contribute greatly to the negative effects for decision making.

Endsley and Bolstad (1994) suggested that in novel situations, decision makers may be forced to use analytic processes that stress limited internal resources. With increasing experience, decision-makers may be able to draw upon mental models and schemata of prototypical situations to provide high levels of situation understanding and hence make good decisions without overloading attention and working memory constraints. ADM is a critical component of pilot proficiency. Current FAA regulations require that decision-making be taught as part of the pilot-training curriculum (FAA, 1991). However, little guidance is provided as to how that might be accomplished, and none is given as to how it might be measured, outside of the practical test.

Researches have suggested that ADM is trainable (Endsley, 1993; Li & Harris, 2005; Orasanu, 1993; Prince & Salas, 1998). Buch and Diehl (1984) found that judgment training subsequently produced significantly better decisions among Canadian civil aviation pilots. Connolly, Blackwell and Lester (1989) also observed that pilot's decision-making skills could be significantly improved through the use of judgment training materials and simulator practice. Nevertheless, Orasanu (1993) pointed out that there was no evidence to support the development of generic training techniques to improve all-purpose decision making skills. It was suggested that different component skills were involved when making the six different basic types of decisions (go/no go decisions; recognition-primed decisions; response selection decisions; resource management decisions; non-diagnostic procedural decisions; and creative problem-solving). These decision types are described in more detail in the following Method section. As a result it was thought unlikely that any one single training method could improve all aspects of decision-making. There were need different types of training for improving the quality of decision-making.

Ⅲ. Mnemonic-based of Decision Support Systems

There are a number of strategies (often embodied in mnemonics or acronyms) describing the processes and procedures concerned with ADM. The common aim of these techniques is to encourage a systematic approach to decision-making that should be less affected by the human nature and should also reduce the cognitive work for pilots (O'Hare, 2003). Four mnemonic methods are considered in the present study. The SHOR mnemonic (Wohl, 1981) consists of four steps: Stimuli, Hypotheses, Options and Response. It was originally developed for use by U.S. Air Force tactical command and control, where decisions were required under high pressure and severe time constraints. In this situation, decisions require near-real-time reactions involving threat warning, task rescheduling and other types of dynamic modification. The SHOR methodology is basically an extension of the stimulus-response paradigm of classical behavioural psychology developed to deal with two aspects of uncertainty in the decision-making process, information input uncertainty followed by the evaluation of the consequences of

actions, which creates the requirement for option generation and evaluation. The PASS methodology was originally developed by Delta Air Lines to train pilots as part of a CRM training program. It consists of four steps: Problem identification (define/redefine problems); Acquire information (seek more information); Survey strategy (survey/resurvey strategies); Select strategy (Maher, 1989). After the selection of a solution strategy, if the problem is not solved, then the pilot should re-enter the problem solving loop once more.

The FORDEC mnemonic was also developed in the civil aviation domain. It was a product of the Lufthansa CRM-course. It comprises of six steps: Facts, Options, Risks and Benefits, Decision, Execution, Check (Hörmann, 1995). It incorporates components addressing gathering data (situation assessment); analysis of risk and benefits (including the effects of time pressure, continually changing conditions and distraction) and having incomplete information. Each step of the FORDEC process uses a guiding question to focus the pilot's attention on a sequence of essential steps for effective decision making. DESIDE (Murray, 1997) was developed on a sample of South African pilots and also comprises of six steps: Detect, Estimate, Set safety objectives, Identify, Do, Evaluate. The DESIDE method is a practical application to aid in-flight decisions adapted from conflict-theory model of Janis and Mann (1977).

Li and Harris (2005) undertook a study to identify the best ADM mnemonic-based methods for training military pilots decision-making in a tactical environment. SHOR was rated as being the best ADM mnemonic in time-limited and critical, urgent situations. DESIDE was regarded as superior for knowledge-based decisions which required more comprehensive considerations but also had more time available to do so. These were subsequently developed into a short ADM training course delivered to cadet pilots in the Republic of China Air Force Tactical Training Wing (Li & Harris, 2008). The pilots who received ADM training exhibited superior decision-making skills during a series of emergency situations presented in a full-flight simulator in terms of situation assessment and risk management, however this was usually at the expense of speed of response.

Pilots of commercial airliners face considerably different decision making challenges than those encountered by pilots of high-performance, single-seat military fighter aircraft. While these aircraft are of lower performance, pilots encounter issues such as those associated with coordinating with a range of national authorities when flying into international airspace; multiple aircraft from other airlines and operating in a multi-crew context. This study was designed to evaluate the potential effectiveness of four aeronautical decision-making mnemonics (SHOR, PASS, FORDEC and DESIDE) in each of the six basic decision –making situations (described by Orasanu, 1993) for use as a basis of a commercial pilot ADM training programme.

Ⅳ. Method

4.1. Participants

One hundred and fifty-seven pilots participated in this research (57 Captains; 99 First Officers; missing data in one case). All participants were type-rated Boeing 747-400 pilots from an International airline. All but five pilots were male.

4.2. Rule-based decision training mnemonics

Previous research demonstrated a belief rule-based decision support system has provided more reliable and informative performance after training (Kong, et al., 2012). This research applied four decision-making mnemonics of rule-based methods that could potentially form the basis of an ADM training program were selected from

a review of the literature by a number of subject matter experts, comprising an airline CRM (Crew Resource Management) director; Boeing 747-400 chief pilots; a Boeing 747-400 training director; an aviation psychologist, and an aviation human factors specialist. The methods finally selected (and previously described briefly in the Mnemonic-based of Decision Support Systems) were:

- SHOR: Stimuli, Hypotheses, Options, Response (Wohl, 1981)
- PASS: Problem identification, Acquire information, Survey strategy, Select strategy (Maher, 1989);
- FORDEC: Facts, Options, Risks & Benefits, Decision, Execution, Check (Hörmann, 1995);
- DESIDE: Detect, Estimate, Set safety objectives, Identify, Do, Evaluate (Murray, 1997).

4.3. The development of scenarios

To develop scenarios for assessing the effectiveness of the ADM mnemonics which corresponded to Orasanu's (1993) six generic decision making categories, six focus groups were conducted, one for each category of scenario. Each focus group comprised two human factors specialists, three senior Boeing 747 instructor pilots and the Director of Crew Resource Management Division of the participating airlines. The purpose of these focus groups was to ensure enough detailed information was contained within each scenario for pilots to make an informed decision and hence to evaluate the applicability for each of the four ADM mnemonics. The scenarios were all developed from actual incidents and accidents. They were as follows.

4.3.1 Go/no go decisions (these decisions are made under severe time pressure and involve considerable risk; the amount of thinking should be minimal): Boeing 747-400 with take-off weight 833,000 pounds. The warning light of 4L door suddenly illuminates while the aircraft is taking off with an indicated air speed of 120 kts.

4.3.2 Recognition-primed decision (there is a need to recognize situational patterns that serve as input to condition-action rules, but in these cases the decision maker must also learn the response side of the rule and its link to that condition): Boeing 747-400 with landing weight 533,000 pounds. The visibility is 3,000 meters, cloud base 500 feet. Autopilot is engaged during the instrument approach, ILS signal is suffering interference and glide slope indication is fluctuating.

4.3.3. Response selection decision (a single option must be selected from a set; crews must recognize multiple options and evaluate them in terms of how well they satisfy the goals and meet constraints): Boeing 747-400 departs from Hong Kong to Taipei with landing weight 533,000 pounds. The ATC clearance "Direct to TONGA, descend and maintain flight level 290, clear to JAMMY via TONGA 3A RNAV ARRIVAL". When aircraft is 3 miles from TONGA, communication is lost, and there is a failure to contact ATC.

4.3.4 Resource management decisions (the relative priorities of critical tasks must be part of a pilot's basic knowledge. Relevant skills to this type of decision include estimation of time required to complete various tasks, knowledge of interdependencies among tasks, and scheduling strategies): Boeing 747-400 with landing weight 533,000 pounds. ATC clearance "Direct to TONGA"; descend and maintain 11,000 feet; clear to JAMMY via "TONGA 3A RNAV ARRIVAL". Three miles before BRAVO, the Captain (PF) is suddenly incapacitated and twice and provides no response to standard CALL OUTs.

4.3.5. Non-diagnostic procedural decisions (these involve responding to a number of cues defining a situation that fall outside

the experience or training of the pilot; the situation has no prescribed response, the nature of this problem is unclear and many different types of ambiguous cues may also signal potentially dangerous conditions): Boeing 747-400 - take-off weight 833,000 pounds; 22:30 local time. When climbing to 1,000 feet with Thrust Reduced to CLB, the aircraft suddenly begins to vibrate significantly. PM observes No.1 ENG vibration indication is abnormal, although other ENG indications are normal. By this time the aircraft has to transition through a cloud area with light turbulence. It is difficult to judge whether the vibration is caused by ENG or turbulence; it is also was unclear whether to continue to destination airport or return to base.

4.3.6. Creative problem-solving (these decisions are the most complex, as they involve both a diagnosis to determine the nature of the situation and response generation. Pilots must determine what their goals are, develop a plan and candidate strategies, and evaluate these strategies and actions based on projections of the potential outcomes): Boeing 747-400 with take-off weight 833,000 pounds. During the climb through 1,000 feet after departure, the fire warning system of No.4 ENG activates, 10 seconds later, the aircraft begins to vibrate heavily and a big "BANG" is heard. The relevant No.4 ENG system fails totally and the fire warning disappears.

4.4. Development of evaluation instrument

To develop a rating instrument for the evaluation of the effectiveness of the SHOR, PASS, FORDEC and DESIDE ADM mnemonics in the in-flight scenarios previously described, six further focus groups were formed, one for each scenario. Each focus group comprised two human factors specialists and three Boeing 747 instructor pilots. The scenarios were analysed by the group members using all four mnemonic methods. This process provided the material for the construction of a self-completion rating form to evaluate the suitability of the ADM mnemonics by a large sample of Boeing 747 crew members.

The narrative responses describing the decision-making process was evaluated using the criteria of 'situation assessment'; 'risk management'; 'response time' and also overall 'applicability'. These dimensions were derived from the earlier studies (Li & Harris, 2005, 2008) which selected the most appropriate ADM training mnemonic methods for fighter pilots and evaluated their efficacy. However, good decisions can lead to bad outcomes (and vice versa) especially when operating in a probabilistic environment, such as aviation. All in-flight decisions are made under some degree of uncertainty. Decision makers cannot infallibly be graded by their results (Brown, Kahr, & Peterson, 1974). Good decisions do not necessarily guarantee a good outcome. Evaluating a decision as good (or not) also depends on the stakes and the processes employed. As a result, the mnemonic methods were evaluation of the effectiveness of the decision-making adjuncts of situation assessment, risk management and response time.

Each mnemonic method was evaluated for its suitability using a structured self-completion rating form using a nine-point Likert-type scale (with a high score of 9 and a low score of 1). Particular care was taken with the 'response time' item where a high score was used to indicate that it was an efficient (i.e., fast) rating. This allowed the scales to be summed meaningfully into a composite score. Further space was allowed for respondents to add qualitative comments justifying their reasons for the ratings awarded.

4.5. Administration of evaluation instrument

The ADM rating instruments were distributed to all pilots of Boeing 747 fleet of the participating airlines. Supporting

material was also provided describing the requirements of each step in the application of each of the mnemonic methods.

As a result of the length of the scenarios and the number of ratings required, each participant only evaluated the ADM mnemonic techniques in three scenarios. The first set of three scenarios consisted of go/no go decisions, response selection decisions, and non-diagnostic procedural decisions; the second set of three scenarios consisted of recognition-primed decisions, resource management decisions and creative problem-solving decisions.
Completed instruments were returned to the Crew Resource Management Division.

V. Results

In total, evaluations of four ADM mnemonics methods were collected from responses to 1,871 flight scenarios. There were 312 completed rating instruments for the go/no-go decision scenario; 311 for the recognition-primed decision-making scenario; 316 for the response selection decision-making scenario; 310 for the resource management scenario; 312 for the non-diagnostic procedural decision-making scenario, and 310 completed rating forms for the creative problem-solving scenario.

5.1. Treatment of data

For each participant an overall composite score for every mnemonic method in all scenarios was created by summing the scores across the four dimensions of situation assessment; risk management; response time; and applicability, and then dividing the resultant total by four (hence the range for the composite scores was the same as that for the individual component scales). To avoid an inflated probability of obtaining a spuriously significant result as a product of performing multiple statistical tests a Bonferroni adjustment was applied and an value of $p < 0.05$ used (Bland & Altman, 1995).

5.2. Go/no-go decisions

There was no significant difference in the ratings of suitability among the four ADM mnemonics ($F_{3,228} = 2.192$; $p = 0.090$) using the composite scale scores in the go/no-go decision making scenario.

5.3. Recognition-primed decisions

There were significant differences in the ratings of suitability among the four ADM mnemonics in the recognition-primed decision making scenario ($F_{3,228} = 5.223$; $p<0.01$) using the composite scale scores. The highest overall rating of the suitability of the ADM mnemonics was for FORDEC followed by DESIDE, PASS and SHOR (see table 1). Further comparisons using post-hoc tests with a Bonferroni adjustment using an alpha value of $p<0.05$ showed significant differences between FORDEC vs. SHOR and FORDEC vs. PASS.

Table 1 Mean and standard for the overall composite scale scores for each ADM mnemonic method in the recognition-primed scenario.

ADM Method	*M*	*SD*
SHOR (77)	6.43	1.57
PASS (77)	6.62	1.24
FORDEC (77)	6.99	1.30
DESIDE (77)	6.77	1.27

As there were significant differences in the composite scores evaluating the suitability of the ADM decision-making methods, to clarify the contribution of each dimension the results were broken down further (see table 2). There was a significant difference in the rated applicability of the methods on the dimension of Situation Assessment ($F_{3,228} = 4.224$; $p<0.05$). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to be rated significantly superior to both SHOR and DESIDE. There were also a significant differences in the assessments of the risk management capability of each ADM mnemonic method ($F_{3,228} = 10.131$; $p<0.001$). Further

comparisons using post-hoc Bonferroni adjusted tests showed FORDEC, PASS and DESIDE to be significantly higher rated than SHOR. There was a significant difference in the ratings concerned with response time when using the ADM mnemonics ($F_{3,225}$ = 5.109; p<0.01). Further post hoc comparisons showed PASS to be rated significant better (i.e., taking less time) than SHOR, and DESIDE was better rated than FORDEC. Finally, there were significant differences in the rating of applicability between the ADM mnemonics ($F_{3,228}$ = 6.635; p<0.001). Post hoc comparisons showed that PASS and DESIDE were significantly better rated than SHOR.

Table 2 Mean and standard deviations for the individual sub-scale means for situation assessment; risk management; response time and overall applicability scales for each ADM mnemonic method in the recognition-primed scenario.

Sub-Scale Rating	ADM Method (n)	M	SD
Situation	SHOR (77)	6.86	1.79
Assessment	PASS (77)	6.99	1.47
	FORDEC (77)	7.38	1.33
	DESIDE (77)	6.99	1.47
Risk	SHOR (77)	6.31	1.87
Management	PASS (77)	6.99	1.47
	FORDEC (77)	7.29	1.44
	DESIDE (77)	6.99	1.47
Response	SHOR (76)	6.41	1.72
Time	PASS (76)	6.97	1.48
	FORDEC (76)	6.41	1.83
	DESIDE (76)	6.97	1.48
Overall	SHOR (77)	6.14	1.91
Applicability	PASS (77)	6.99	1.47
	FORDEC (77)	6.88	1.72
	DESIDE (77)	6.99	1.47

5.4. Response selection decisions

There were significant differences in the ratings of suitability among the four ADM mnemonics in the response selection decision making scenario ($F_{3,228}$ = 14.637; p<0.001) using the composite scale scores. The highest overall rating of the suitability of the ADM mnemonics was for FORDEC followed by DESIDE, PASS and SHOR (see table 3). Further comparisons using post-hoc tests with a Bonferroni adjustment using an alpha value of p<0.05 showed significant differences between FORDEC vs. SHOR; FORDEC vs. PASS; FORDEC vs. DESIDE and DESIDE vs. SHOR.

Table 3 Mean and standard for the overall composite scale scores for each ADM mnemonic method in the response selection scenario.

ADM Method	M	SD
SHOR (77)	6.59	1.15
PASS (77)	6.82	1.03
FORDEC (77)	7.44	1.11
DESIDE (77)	7.00	1.22

As there were significant differences in the composite scores evaluating the suitability of the ADM decision-making methods, to clarify the contribution of each dimension the results were broken down further (see table 4). There was a significant difference in the rated applicability of the methods on the dimension of Situation Assessment ($F_{3,234}$ = 10.539 p<0.001). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to be rated significantly superior to both SHOR and DESIDE. DESIDE was rated as being significantly superior to SHOR on the dimension of Situation Assessment. There was a significant difference in the ratings of Risk Management among the ADM mnemonics ($F_{3,231}$=14.317; p<0.001). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to have significantly higher ratings in this aspect than SHOR, PASS and DESIDE. PASS was also rated to be significantly superior to SHOR in this respect. There was only one significant difference in the ratings of Response Time between the mnemonic methods ($F_{3,231}$ = 4.541; p<0.001). Post-hoc comparisons showed FORDEC to be better rated than PASS. Finally, were also

significant differences in the rated overall Applicability of the ADM mnemonics ($F_{3,231}$ = 16.621; p<0.001). FORDEC had significantly superior ratings than SHOR, PASS and DESIDE. DESIDE was significantly more highly rated than SHOR.

Table 4 Mean and standard deviations for the individual sub-scale means for situation assessment; risk management; response time and overall applicability scales for each ADM mnemonic method in the response selection scenario.

Sub-Scale Rating	ADM Method (n)	M	SD
Situation	SHOR (79)	6.67	1.41
Assessment	PASS (79)	6.85	1.21
	FORDEC (79)	7.47	1.15
	DESIDE (79)	7.14	1.27
Risk	SHOR (78)	6.44	1.41
Management	PASS (78)	7.01	1.12
	FORDEC (78)	7.49	1.15
	DESIDE (78)	6.90	1.34
Response	SHOR (78)	6.76	1.12
Time	PASS (78)	6.68	1.10
	FORDEC (78)	7.21	1.37
	DESIDE (78)	6.85	1.30
Overall	SHOR (78)	6.51	1.45
Applicability	PASS (78)	6.76	1.23
	FORDEC (78)	7.59	1.20
	DESIDE (78)	7.10	1.30

5.5. Resource management decisions

There was no significant difference in the ratings of suitability among the four ADM mnemonics ($F_{3,228}$ = 2.639; p = 0.051) using the composite scale scores in the go/no-go decision making scenario.

5.6. Non-diagnostic procedural decisions

There were significant differences in the ratings of suitability among the four ADM mnemonics in the non-diagnostic procedural decision making scenario ($F_{3,228}$ =20.494 p<0.001) using the composite scale scores. The highest overall rating of the suitability of the ADM mnemonics was for again for FORDEC followed by DESIDE, PASS and SHOR (see table 5). Further comparisons using post-hoc tests with a Bonferroni adjustment using an alpha value of $p < 0.05$ showed significant differences between FORDEC vs. SHOR; FORDEC vs. PASS; FORDEC vs. DESIDE; DESIDE vs. SHOR and DESIDE vs. PASS.

Table 5 Mean and standard deviations for the overall composite scale scores for each ADM mnemonic method in the non-diagnostic procedural scenario.

ADM Method	M	SD
SHOR (77)	6.49	1.33
PASS (77)	6.73	1.12
FORDEC (77)	7.50	1.16
DESIDE (77)	7.09	1.10

As there were significant differences in the composite scores evaluating the suitability of the ADM decision-making methods, to clarify the contribution of each dimension the results were broken down further (see table 6). There was a significant difference in the rated applicability of the ADM methods on the dimension of Situation Assessment ($F_{3,228}$=21.654; p<0.001). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to have a significantly higher score than both SHOR and DESIDE, and DESIDE to have a superior score than both SHOR and PASS. There was a significant difference in the ratings of Risk Management among the ADM mnemonics ($F_{3,228}$ = 15.830; p<0.001). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to have significant higher scores than both SHOR and PASS, and DESIDE to be better rated than SHOR on this sub-scale. There were also significant differences in the ratings of response time between the ADM mnemonics ($F_{3,225}$ = 9.350; p<0.001). Post-hoc comparisons showed FORDEC had significantly higher score than SHOR and PASS, and DESIDE was significantly better than SHOR in this respect. Finally in this scenario, there were also significant differences in the rated overall Applicability of the ADM

mnemonics ($F_{3,234}$ = 17.250; p<0.001). Post-hoc comparisons showed FORDEC to be better rated overall than SHOR, PASS and DESIDE. DESIDE received significantly better ratings than SHOR.

Table 6 Mean and standard deviations for the individual sub-scale means for situation assessment; risk management; response time and overall applicability scales for each ADM mnemonic method in the non-diagnostic procedural scenario.

Sub-Scale Rating	ADM Method (n)	M	SD
Situation Assessment	SHOR (77)	6.57	1.52
	PASS (77)	6.78	1.19
	FORDEC (77)	7.57	1.12
	DESIDE (77)	7.16	1.17
Risk Management	SHOR (77)	6.45	1.49
	PASS (77)	6.81	1.21
	FORDEC (77)	7.57	1.33
	DESIDE (77)	7.12	1.18
Response Time	SHOR (76)	6.45	1.40
	PASS (76)	6.63	1.19
	FORDEC (76)	7.25	1.43
	DESIDE (76)	7.04	1.32
Overall Applicability	SHOR (79)	6.48	1.39
	PASS (79)	6.69	1.34
	FORDEC (79)	7.60	1.24
	DESIDE (79)	7.05	1.20

5.7. Problem-solving decisions

There were significant differences in the ratings of suitability among the four ADM mnemonics in the non-diagnostic procedural decision making scenario ($F_{3,225}$ = 3.379; p<0.05) using the composite scale scores. The highest overall score for the suitability of the ADM mnemonics was for FORDEC followed by DESIDE, PASS and SHOR. Further comparisons using post-hoc tests with Bonferroni adjustment showed significant differences only between FORDEC vs. PASS (see table 7).

Table 7 Mean and standard deviations for the overall composite scale scores for each ADM mnemonic method in the problem solving scenario.

ADM Method	M	SD
SHOR (77)	6.79	1.46
PASS (77)	6.80	1.15
FORDEC (77)	7.21	1.34
DESIDE (77)	6.91	1.19

As there were significant differences in the composite scores evaluating the suitability of the ADM decision-making methods, to clarify the contribution of each dimension the results were broken down further (see table 8). There was a significant difference on the dimension of Situation Assessment between the ADM mnemonics ($F_{3,225}$ = 3.390; p<0.05). Further comparisons using post-hoc Bonferroni adjusted tests showed FORDEC to have a significantly higher score than PASS. There was also a significant difference in Risk Management between the techniques ($F_{3,225}$ = 6.679; p<0.001). Post hoc tests showed FORDEC to be significantly better rated than SHOR, PASS and DESIDE. There was, however, no significant difference in the rated response time among the mnemonics ($F_{3,225}$=1.227; p = 0.301). Finally, while there was an overall significant difference in the rated Overall Applicability of the ADM mnemonics ($F_{3,222}$ = 3.129; p<0.05), further post hoc comparisons showed no significant differences between the techniques.

Table 8 Mean and standard for the individual sub-scale means for situation assessment; risk management; response time and overall applicability scales for each ADM mnemonic method in the problem solving scenario.

Sub-Scale Rating	ADM Method (n)	M	SD
Situation Assessment	SHOR (76)	6.97	1.58
	PASS (76)	7.00	1.34
	FORDEC (76)	7.38	1.43
	DESIDE (76)	7.26	1.32
Risk Management	SHOR (76)	6.57	1.58
	PASS (76)	6.87	1.30
	FORDEC (76)	7.34	1.38
	DESIDE (76)	6.89	1.25
Response	SHOR (76)	6.92	1.64

Time	PASS (76)	6.74	1.30
	FORDEC (76)	6.93	1.54
	DESIDE (76)	6.63	1.39
Overall	SHOR (75)	6.68	1.67
Applicability	PASS (75)	6.60	1.26
	FORDEC (75)	7.15	1.71
	DESIDE (75)	6.85	1.46

VI. Discussion

In commercial flight operations, pilots are confronted with many problems occurring in continually changing situations that have the potential to lead to human error and occasionally accidents. Analysis of commercial aviation mishap data have suggested that decision errors can account for between 40%, and as much as 70% of accidents (Diehl, 1991; Li, et al., 2008). However, it has also been shown that training using mnemonic-based methods can significantly improve pilot decision making (Li & Harris, 2008) when the most appropriate method is applied for a particular category of scenario.

This study surveyed a large sample of qualified airline pilots flying the Boeing 747-400 to evaluate the potential effectiveness of four aeronautical decision-making mnemonic methods (SHOR, PASS, FORDEC and DESIDE) in each of the six basic decision-making situations (described by Orasanu, 1993) for use as the basis of a commercial pilot ADM training programme. The work was based upon an earlier study undertaken by Li and Harris (2005) who used a similar approach to identify the best ADM mnemonic-based methods for training military pilots' tactical decision-making.

In the present study FORDEC (Hörmann, 1995) was consistently evaluated as being the overall highest-rated (in terms of its composite scale score) across all six different decision-making scenarios (see summary of results in table 9). DESIDE (Murray, 1997) was generally rated as the second-best method, followed by PASS (Maher, 1989) and finally SHOR (Wohl, 1981).

The data also indicated that FORDEC was rated highly for the individual elements underpinning good decision-making, such as situation assessment and risk management (see tables 2, 4, 6 and 8). Respondents thought that the technique was comprehensive and thorough; clear about how to identify the safest actions; and it also had a logical order and was easy to remember. Furthermore, it matched the general format of a typical pre-flight briefing. However, it was also assessed as requiring more time to perform the required analysis and produce a final decision (see tables 2, 4, 6 and 8). FORDEC's characteristics were thought to be best suited when dealing with non-urgent situations. Many pilots advised practicing FORDEC in the simulator was extremely important before attempting to apply it in a real life situation. It was also rated as the best ADM mnemonic–based decision making method for promoting crew coordination, as would be expected of a methodology originally developed to encourage good CRM.

Table 9 Summary of rankings for the four ADM mnemonic methods across the six decision making scenarios (lower numbers are better-rated)

	Scenarios					
Mnemonic Method	Go/No go decision	Recognition-primed decision	Response selection decision	Resource management decision	Non-diagnostic procedural decision	Creative problem-solving
SHOR	2	4	4	3	4	4
PASS	4	3	3	4	3	3
FORDEC	1	1	1	1	1	1
DESIDE	3	2	2	2	2	2

DESIDE was also regarded by respondents as being comprehensive but was another time consuming method. PASS was also in accord with the airlines pilot's training guidelines as it had clear and specific procedures to follow, but was relatively poorly rated compared to the

other methods despite it also being developed for application in an airline (Delta). The qualitative data suggested that SHOR was regarded by pilots as a method for promoting quick decision-making in urgent situations with a logical order for flight operations safely.

The pattern of results in this instance show considerable differences from those obtained by Li and Harris (2005) which looked at decision making in a military fast-jet environment. In this earlier study SHOR was assessed as the best ADM training mnemonic method in scenarios where fast responses were required. PASS was also highly rated in such scenarios. It should be noted that SHOR was developed for use by the U.S. Air Force for tactical command and control situations, where decisions are often made under time pressure and within severe time constraints. In this earlier study instructors regarded them as the best methods for rule-based decisions (condition-action rules). In the earlier Li and Harris study DESIDE was evaluated as being the best ADM method to use when making response selection decisions and non-diagnostic procedural decisions. These are non-emergent situations and have no immediately dangerous threats hence pilots have time to consider their actions more extensively. FORDEC was only highly rated for making knowledge-based decisions (for well-defined problems) which included resource management decisions. In contrast, in this study FORDEC was universally the most highly rated ADM technique.

The requirements for military fast jet flying and operating as part of a crew in a modern commercial airliner are quite different. Many military fast jets are single seat aircraft (although modified CRM concepts are still taught as aircraft tend to operate in a pair or as part of a four-ship formation) however they are of much higher performance and the operating environment is considerably more uncertain and prone to rapid changes. In contrast, airline operations are more predictable and highly proceduralied. Pilots fly as part of a crew which extends beyond the flight deck door. While things can occasionally go wrong quite quickly, the tempo of operations is considerably slower than their military counterparts. As a result, the majority of situations faced by airline pilots allow the time for a more considered response, hence the preference for FORDEC, as opposed to the less comprehensive but faster SHOR and PASS decision-making mnemonic methods.

Kaempf and Orasanu's (1997) results demonstrated that under conditions of time pressure, decision makers need help to determine what is occurring in the environment around them. This suggests that if decision-making is required in such circumstances, practice should include task performance under those conditions (see Connolly, et al., 1989). Decision aids and training based upon mnemonic methods can provide decision makers on the flight deck with the tools and skills necessary to make quick, accurate situation assessments and employ appropriate risk management strategies (Li & Harris, 2008).

The ability to make decisions in the air has often been regarded as by-product of flying experience rather than training. However, the results from decision-making training programmes suggest otherwise (Buch & Diehl, 1984; Endsley, 1993; Orasanu, 1993; Prince & Salas, 1998) and the results of this research suggest that FORDEC can provide a suitable basis for such training. Furthermore, in the US the FAA requires that decision-making be taught as part of the pilot-training curriculum but does not specify how this should be undertaken (FAA, 1991). The FORDEC method was rated as being applicable for all six basic types of decision described by Orasanu (1993). The principal limitation of the present study was that it only elicited pilots' opinions about the efficacy of these decision-making techniques. As a result, research needs to be undertaken to produce

empirical performance data to establish if training in the use of ADM mnemonic-based methods such as FORDEC does actually improve pilots' in-flight decision-making.

Ⅶ. Conclusions

The data in the present study were gathered from airline pilots evaluating the effectiveness of ADM mnemonics based on scenarios of flight operations. The results suggest that FORDEC is the best ADM mnemonic decision-making aid for use across all in-flight situations. As a result the FORDEC mnemonic forms a suitable basis for airline pilot decision-making training. However, it is a compromise of more response time for FORDEC, as in-depth considerations to deal with uncertain and time-limited situations more time for cognitive processing the information. To ensure the effectiveness of such decision-making training, it will be necessary to deliver instruction using the FORDEC mnemonic-based method and practice application of the approach in a flight simulator.

References

[1] Bland, J.M., & Altman, D.G. (1995). Multiple significance tests: The Bonferroni method. *BMJ*, 310, 170.

[2] Brown, R. V., Kahr, A. S., & Peterson, C. (1974). *Decision analysis for the manager*. New York: Holt, Rinehart and Winston.

[3] Buch, G., & Diehl, A. (1984). An investigation of the effectiveness of pilot judgment training. Human Factors: *The Journal of the Human Factors and Ergonomics Society*, 26, 557-564.

[4] Connolly, T. J., Blackwell, B. B., & Lester, L. F. (1989). A simulator-based approach to training in aeronautical decision making. *Aviation, space, and environmental medicine*, 60, 50-52.

[5] Diehl, A. (1991). The effectiveness of training programs for preventing aircrew error. In proceedings of the Sixth *International Symposium on Aviation Psychology* (pp. 640-655). Columbus, OH.

[6] Dreiseitl, S., & Binder, M. (2005). Do physicians value decision support? A look at the effect of decision support systems on physician opinion. *Artificial intelligence in medicine*, 33, 25-30.

[7] Endsley, M. R. (1993). A survey of situation awareness requirements in air-to-air combat fighters. *The International Journal of Aviation Psychology*, 3, 157-168.

[8] Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *The International Journal of Aviation Psychology*, 4, 241-264.

[9] Federal Aviation Administration. (1991). *Aeronautical Decision-making* (Advisory Circular 60-22). In Washington, DC: US Department of Transport.

[10] Hörmann, H. J. (1995). FORDEC: A perspective model for aeronautical decision making. In R. Fuller, N. Johnston & N. McDonald (Eds.), *Human factors in aviation operations* (pp. 17-23). Aldershot, England: Ashgate.

[11] Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York, NY: Free Press.

[12] Jensen, R. S., & Benel, R. A. (1977). Judgment evaluation and instruction in civil pilot training. In. Washington, DC: *Federal Aviation Administration*.

[13] Kaempf, G. L., & Orsanu, J. (1997). Current and future applications of naturalistic decision making in aviation. *In Naturalistic decision making* (pp. 81-90). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[14] Kong, G., Xu, D.-L., Body, R., Yang,

J.-B., Mackway-Jones, K., & Carley, S. (2012). A belief rule-based decision support system for clinical risk assessment of cardiac chest pain. *European Journal of Operational Research*, 219, 564-573. DOI: 10.1016/j.ejor.2011.10.044

[15] Li, W. C., & Harris, D. (2005). Aeronautical decision making: Instructor-pilot evaluation of five mnemonic methods. *Aviation, space, and environmental medicine*, 76, 1156-1161.

[16] Li, W. C., & Harris, D. (2008). The evaluation of the effect of a short aeronautical decision-making training program for military pilots. *The International Journal of Aviation Psychology*, 18, 135-152. DOI: 10.1080/10508410801926715

[17] Li, W. C., Harris, D., & Yu, C. S. (2008). Routes to failure: Analysis of 41 civil aviation accidents from the Republic of China using the human factors analysis and classification system. *Accident Analysis & Prevention*, 40, 426-434. DOI: 10.1016/j.aap.2007.07.011

[18] Liao, S.-H. (2000). Case-based decision support system: architecture for simulating military command and control. *European Journal of Operational Research*, 123, 558-567. DOI: 10.1016/S0377-2217(99)00109-5

[19] Maher, J. (1989). Beyond CRM to decisional heuristics: An airline generated model to examine accidents and incidents caused by crew errors in deciding. *In proceedings of the Fifth International Symposium on Aviation Psychology* (pp. 439-444). Columbus, OH.

[20] McKinney, E. H. (1993). Flight leads and crisis decision-making. *Aviation, space, and environmental medicine*, 64, 359-362.

[21] Murray, S. R. (1997). Deliberate decision making by aircraft pilots: A simple reminder to avoid decision making under panic. *The International Journal of Aviation Psychology*, 7, 83-100. DOI: 10.1207/s15327108ijap0701_5

[22] O'Hare, D. (2003). Aeronautical decision making: Metaphors, models, and methods. In P. S. Tsang & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 201-237). Mahwah NJ: Lawrence Erlbaum.

[23] Orasanu, J. (1993). Decision-making in the cockpit. In E. L. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit Resource Management* (pp. 137-172). San Diego, CA: Academic Press.

[24] Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu & R. Z. Calderwood, C. E. (Eds.), *Decision Making in Action: Models and Methods*. (pp. 3-20). Norwood, NJ: Ablex Publishing Corporation.

[25] Power, D. (2002). *Decision support systems: concepts and resources for managers*. Westport, Connecticut: Quorum Books.

[26] Prince, C., & Salas, E. (1998). Situation assessment for routine flight and decision making. *International Journal of Cognitive Ergonomics*, 1, 315-324.

[27] Shappell, S., & Wiegmann, D. (2004). HFACS analysis of military and civilian aviation accidents: A North American comparison. *In proceedings of the International Society of Air Safety Investigators*. Queensland, Australia.

[28] Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Aldershot, England: Ashgate Publishing.

[29] Wohl, J. G. (1981). Force management decision requirements for Air Force tactical command and control. *Systems, Man and*

Cybernetics, IEEE Transactions on, 11, 618-639. DOI: 10.1109/TSMC.1981.4308760

2014-06-01

The application of aeronautical decision-making support systems for improving pilots' performance in flight operations

Li, Wen-Chin

Aeronautical and Astronautical Society of the Republic of China

Wen-Chin Li, Lun-Wen Li, Don Harris and Yueh-Ling Hsu. The application of aeronautical decision-making support systems for improving pilots' performance in flight operations. Journal of Aeronautics, Astronautics and Aviation. 2014, Vol. 46, Iss. 2, pp114-123

http://dx.doi.org/10.6125/14-0324-789